Seqlist.txt

SEQUENCE LISTING

<110> BRETT, Sara
      HAMPLIN, Paul Andrew
      OGILVIE, Louise

<120> Vaccine Against HCV

<130> PG5029

<140> 10/535,047
<141> 2005-05-13

<150> PCT/EP03/12830
<151> 2003-11-13

<150> 0226722.7
<151> 2002-11-15

<160> 24

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 60
<212> DNA
<213> Hepatitis C virus

<400> 1
gaattcgcgg ccgccatgag caccaacccc aagccccagc gcaagaccaa gcggaacacc 60

<210> 2
<211> 59
<212> DNA
<213> Hepatitis C virus

<400> 2
gaattcggat cctcatgcgc tagcggggat ggtgaggcag ctcagcagcg ccagcagga 59

<210> 3
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 3
gaattcgcgg ccgccatggc ccccatcacc gcctacagcc agcagacccg gggac   55

<210> 4
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 4
gaattcggat cctcaggtga ccacctccag gtcagcggac atgcacgcca tgatg   55

<210> 5
<211> 46
<212> DNA
<213> Hepatitis C virus

<400> 5
gaattcgcgg ccgccatgtt ttgggccaag catatgtgga acttca   46

<210> 6
<211> 46
<212> DNA

<213> Hepatitis C virus

<400> 6
gaattcggat cctcagcaag gggtggagca gtcctcgttg atccac   46

<210> 7
<211> 49
<212> DNA
<213> Hepatitis C virus

<400> 7
gaattcgcgg ccgccatgtc catgtcctac acctggaccg gcgccctga   49

<210> 8
<211> 49
<212> DNA
<213> Hepatitis C virus

<400> 8
gaattcggat cctcagcggt tgggcagcag gtagatgccg actccgacg   49

<210> 9
<211> 191
<212> PRT
<213> Hepatitis C virus·

<400> 9
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
1   5   10   15
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20   25   30
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
            35   40   45
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
    50   55   60
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
65   70   75   80
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
                85   90   95
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
                100   105   110
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
            115   120   125
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
            130   135   140
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145   150   155   160
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
                165   170   175
Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala
                180   185   190

<210> 10
<211> 632
<212> PRT
<213> Hepatitis C virus

<400> 10
Met Ala Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu Leu Gly
1   5   10   15
Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly
            20   25   30
Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr Cys
            35   40   45
Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys Thr
    50   55   60
Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val Asp

```
65                      70                      75                      80
Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Met Thr
                    85                      90                      95
Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala
                100                     105                     110
Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu
            115                     120                     125
Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Val Gly Gly Pro Leu
        130                     135                     140
Leu Cys Pro Ser Gly His Val Val Gly Ile Phe Arg Ala Ala Val Cys
145                     150                     155                     160
Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser Met
                165                     170                     175
Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro
                180                     185                     190
Ala Val Pro Gln Thr Phe Gln Val Ala His Leu His Ala Pro Thr Gly
        195                     200                     205
Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr
        210                     215                     220
Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly
225                     230                     235                     240
Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly
                245                     250                     255
Val Arg Thr Ile Thr Thr Gly Ala Pro Ile Thr Tyr Ser Thr Tyr Gly
                260                     265                     270
Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile
            275                     280                     285
Ile Cys Gln Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu Gly Ile
            290                     295                     300
Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val
305                     310                     315                     320
Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn
                325                     330                     335
Ile Glu Glu Val Ala Leu Ser Asn Asn Gly Glu Ile Pro Phe Tyr Gly
                340                     345                     350
Lys Ala Ile Pro Ile Glu Ala Ile Lys Gly Gly Arg His Leu Ile Phe
            355                     360                     365
Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Ser Gly
            370                     375                     380
Leu Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val
385                     390                     395                     400
Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
                405                     410                     415
Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys
                420                     425                     430
Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu
            435                     440                     445
Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly
450                     455                     460
Arg Thr Gly Arg Gly Arg Ser Gly Ile Tyr Arg Phe Val Thr Pro Gly
465                     470                     475                     480
Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr
                485                     490                     495
Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser Val
                500                     505                     510
Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
        515                     520                     525
His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile Asp
        530                     535                     540
Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro Tyr
545                     550                     555                     560
Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro
                565                     570                     575
Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr
                580                     585                     590
Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn
        595                     600                     605
Glu Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys Met
                            Page 3
```

```
         610                    615                    620
Ser Ala Asp Leu Glu Val Val Thr
625                    630


<210> 11
<211> 214
<212> PRT
<213> Hepatitis C virus

<400> 11
Met Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr
1               5                   10                  15
Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu
            20                  25                  30
Met Ala Phe Thr Ala Ser Ile Thr Ser Pro Leu Thr Thr Gln Asn Thr
        35                  40                  45
Leu Leu Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Pro
    50                  55                  60
Pro Ser Ala Ala Ser Ala Phe Val Gly Ala Gly Ile Ala Gly Ala Ala
65                  70                  75                  80
Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly
            85                  90                  95
Tyr Gly Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Val Met Ser
            100                 105                 110
Gly Glu Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile
        115                 120                 125
Leu Ser Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu
    130                 135                 140
Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg
145                 150                 155                 160
Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr
                165                 170                 175
Val Pro Glu Ser Asp Ala Ala Ala Arg Val Thr Gln Ile Leu Ser Ser
            180                 185                 190
Leu Thr Ile Thr Gln Leu Leu Lys Arg Leu His Gln Trp Ile Asn Glu
            195                 200                 205
Asp Cys Ser Thr Pro Cys
        210


<210> 12
<211> 592
<212> PRT
<213> Hepatitis C virus

<400> 12
Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
1               5                   10                  15
Ala Glu Glu Ser Lys Leu Pro Ile Asn Pro Leu Ser Asn Ser Leu Leu
            20                  25                  30
Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Ser Leu
        35                  40                  45
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
    50                  55                  60
Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys
65                  70                  75                  80
Ala Lys Leu Leu Ser Ile Glu Glu Ala Cys Lys Leu Thr Pro Pro His
                85                  90                  95
Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
            100                 105                 110
Ser Ser Arg Ala Val Asn His Ile Arg Ser Val Trp Glu Asp Leu Leu
        115                 120                 125
Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Ser Glu
    130                 135                 140
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
145                 150                 155                 160
Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
                        Page 4
```

```
                        165                   170                   175
Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr
                180                   185                   190
Gly Phe Gln Tyr Ser Pro Lys Gln Arg Val Glu Phe Leu Val Asn Thr
        195                   200                   205
Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ser Tyr Gly Thr Arg Cys
    210                   215                   220
Phe Gly Ser Thr Val Thr Glu Ser Asp Ile Arg Val Glu Glu Ser Ile
225                   230                   235                   240
Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser
                245                   250                   255
Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
                260                   265                   270
Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
        275                   280                   285
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Thr Ala Ala Cys
    290                   295                   300
Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp
305                   310                   315                   320
Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ala
                325                   330                   335
Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
                340                   345                   350
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
        355                   360                   365
Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
    370                   375                   380
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
385                   390                   395                   400
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
                405                   410                   415
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
                420                   425                   430
Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile
        435                   440                   445
Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
    450                   455                   460
Glu Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
465                   470                   475                   480
Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
                485                   490                   495
Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Lys Leu
                500                   505                   510
Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Arg Tyr Leu Phe Asn
        515                   520                   525
Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
        530                   535                   540
Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
545                   550                   555                   560
Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Pro Leu Cys
                565                   570                   575
Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
                580                   585                   590
```

<210> 13
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 13

```
Pro Arg Phe Gly Lys Ala Ile Pro Ile Glu Ala Ile Lys Gly Gly
 1               5                   10                  15
```

<210> 14
<211> 15
<212> PRT

<213> Hepatitis C virus

<400> 14
Tyr Arg Leu Gly Ala Val Gln Asn Glu Val Ile Leu Thr His Pro
1   5   10   15

<210> 15
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 15
Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala
1   5   10   15

<210> 16
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 16
Ala Ala Ala Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser
1   5   10   15

<210> 17
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 17
Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala
1   5   10   15

<210> 18
<211> 15
<212> PRT
<213> Hepatitis C virus

<400> 18
Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Trp Tyr
1   5   10   15

<210> 19
<211> 9595
<212> DNA
<213> Hepatitis C virus

<400> 19
gccagccccc tgatgggggc gacactccac catgaatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
cccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacc gccgcccaca ggacgtcaag ttcccgggcg 420
gtggtcagat cgttggtgga gtttacctgt tgccgcgcag gggccccagg ttgggtgtgc 480
gcgcgactag gaaggcttcc gagcggtcgc aacctcgtgg aaggcgacaa cctatcccaa 540
aggctcgccg acccgagggc agggcctggg ctcagcccgg gtacccttgg cccctctatg 600
gcaatgaggg cctggggtgg gcaggatggc tcctgtcacc ccgcggctcc cggcctagtt 660
ggggccccac ggacccccgg cgtaggtcgc gtaacttggg taaggtcatc gatacccttg 720
catgcggctt cgccgatctc atggggtaca ttccgctcgt cggcgccccc ctaggggcg 780
ctgccagggc cttggcacac ggtgtccggg ttctggagga cggcgtgaac tatgcaacag 840
ggaacttgcc cggttgctct ttctctatct tcctcttggc tctgctgtcc tgtttgacca 900

```
tcccagcttc cgcttatgaa gtgcgcaacg tgtccgggat ataccatgtc acgaacgact 960
gctccaactc aagcattgtg tatgaggcag cggacgtgat catgcatact cccgggtgcg 1020
tgccctgtgt tcaggagggt aacagctccc gttgctgggt agcgctcact cccacgctcg 1080
cggccaggaa tgccagcgtc cccactacga caatacgacg ccacgtcgac ttgctcgttg 1140
ggacggctgc tttctgctcc gctatgtacg tgggggatct ctgcggatct attttcctcg 1200
tctcccagct gttcaccttc tcgcctcgcc ggcatgagac agtgcaggac tgcaactgct 1260
caatctatcc cggccatgta tcaggtcacc gcatggcttg ggatatgatg atgaactggt 1320
cacctacaac agccctagtg gtgtcgcagt tgctccggat cccacaagct gtcgtggaca 1380
tggtggcggg ggcccactgg ggagtcctgg cgggccttgc ctactattcc atggtaggga 1440
actgggctaa ggttctgatt gtggcgctac tctttgccgg cgttgacggg gagacccaca 1500
cgacggggag ggtggccggc cacaccacct ccgggttcac gtccctttc tcatctgggg 1560
cgtctcagaa aatccagctt gtgaatacca acggcagctg gcacatcaac aggactgccc 1620
taaattgcaa tgactccctc caaactgggt tctttgccgc gctgtttttac gcacacaagt 1680
tcaactcgtc cgggtgcccg gagcgcatgg ccagctgccg ccccattgac tggttcgccc 1740
aggggtgggg ccccatcacc tatactaagc ctaacagctc ggatcagagg ccttattgct 1800
ggcattacgc gcctcgaccg tgtggtgtcg tacccgcgtc gcaggtgtgt ggtccagtgt 1860
attgtttcac cccaagccct gttgtggtgg ggaccaccga tcgttccggt gtccctacgt 1920
atagctgggg ggagaatgag acagacgtga tgctcctcaa caacacgcgt ccgccacaag 1980
gcaactggtt cggctgtaca tggatgaata gtactgggtt cactaagacg tgcggaggtc 2040
ccccgtgtaa catcgggggg gtcggtaacc gcaccttgat ctgccccacg gactgcttcc 2100
ggaagcaccc cgaggctact tacacaaaat gtggctcggg gccctggttg acacctaggt 2160
gcctagtaga ctacccatac aggcttttggc actacccctg cactctcaat ttttccatct 2220
ttaaggttag gatgtatgtg gggggcgtgg agcacaggct caatgccgca tgcaattgga 2280
ctcgaggaga gcgctgtaac ttggaggaca gggataggtc agaactcagc ccgctgctgc 2340
tgtctacaac agagtggcag atactgccct gtgctttcac caccctaccg gctttatcca 2400
ctggtttgat ccatctccat cagaacatcg tggacgtgca atacctgtac ggtgtagggt 2460
cagcgtttgt ctcctttgca atcaaatggg agtacatcct gttgcttttc cttctcctgg 2520
cagacgcgcg cgtgtgtgcc tgcttgtgga tgatgctgct gatagcccag gctgaggccg 2580
ccttagagaa cttggtggtc ctcaatgcgg cgtccgtggc cggagcgcat ggtattctct 2640
cctttcttgt gttcttctgc gccgcctggt acattaaggg caggctggct cctggggcgg 2700
cgtatgcttt ttatggcgtca tggccgctgc tcctgcctcct actggcgtta ccaccacgag 2760
cttacgcctt ggaccgggag atggctgcat cgtgcggggg tgcggttctt gtaggtctgg 2820
tattcttgac cttgtcacca tactacaaag tgtttctcac taggctcata tggtggttac 2880
aatactttat caccagagcc gaggcgcaca tgcaagtgtg ggtcccccc ctcaacgttc 2940
ggggaggccg cgatgccatc atcctcctca cgtgtgcggt tcatccagag ttaatttttg 3000
acatcaccaa actcctgctc gccatactcg gcccgctcat ggtgctccag gctggcataa 3060
cgagagtgcc gtacttcgtg cgcgctcaag ggctcattcg tgcatgcatg ttagtgcgaa 3120
aagtcgccgg gggtcattat gtccaaatgg tcttcatgaa gctgggcgcg ctgacaggta 3180
cgtacgttta taaccatctt accccactgc gggactgggc ccacgcgggc ctacgagacc 3240
ttgcggtggc ggtagagccc gtcgtcttct ccgccatgga gaccaaggtc atcacctggg 3300
gagcagacac cgctgcgtgt gggacatca tcttgggtct acccgtctcc gcccgaaggg 3360
ggaaggagat attttttggga ccggctgata gtctcgaagg gcaagggtgg cgactccttg 3420
cgcccatcac ggcctactcc caacaaacgc ggggcgtact tggttgcatc atcactagcc 3480
tcacaggccg ggacaagaac caggtcgaag gggaggttca agtggtttct accgcaacac 3540
aatctttcct ggcgacctgc atcaacggcg tgtgctggac tgtctaccat ggcgctggct 3600
cgaagaccct agccggtcca aaaggtccaa tcacccaaat gtacaccaat gtagacctgg 3660
acctcgtcgg ctggcaggcg cccccgggg cgcgctccat gacaccatgc agctgtggca 3720
gctcggacct ttacttggtc acgagacatg ctgatgtcat tccggtgcgc cggcgaggcg 3780
acagcagggg aagtctactc tcccccaggc ccgtctccta cctgaaaggc tcctcgggtg 3840
gtccattgct ttgcccttcg gggcacgtcg tgggcgtctt ccgggctgct gtgtgcaccc 3900
gggggtcgc gaaggcggtg gacttcatac ccgttgagtc tatggaaact accatggit 3960
ctccggtctt cacagacaac tcaacccccc cggctgtacc gcagacattc caagtggcac 4020
atctgcacgc tcctactggc agcggcaaga gcaccaaagt gccggctgcg tatgcagccc 4080
aagggtacaa ggtgctcgtc ctgaacccgt ccgttgccgc caccttaggg tttggggcgt 4140
atatgtccaa ggcacacggt atcgaccta acatcagaac tggggtaagg accattacca 4200
cgggcggctc cattacgtac tccacctatg gcaagttcct tgccgacggt ggctgttctg 4260
ggggcgccta tgacatcata atatgtgatg agtgccactc aactgactcg actaccatct 4320
tgggcatcgg cacagtcctg gaccaagcgg agacggctgg agcgcggctc gtcgtgctcg 4380
ccaccgctac acctccggga tcggttaccg tgccacaccc caatatcgag gaaataggct 4440
tgtccaacaa tggagagatc ccccttctatg gcaaagccat ccccattgag gccatcaagg 4500
ggggaggca tctcatttttc tgccattcca agaagaaatg tgacgagctc gccgcaaagc 4560
tgacaggcct cggactgaac gctgtagcat attaccgggg ccttgatgtg tccgtcatac 4620
cgcctatcgg agacgtcgtt gtcgtggcaa cagacgctct aatgacgggt ttcaccggcg 4680
attttgactc agtgatcgac tgcaatacat gtgtcaccca gacagtcgac ttcagcttgg 4740
atccccactt caccattgag acgacgaccg tgccccaaga cgcggtgtcg cgctcgcaac 4800
ggcgaggtag aactggcagg ggtaggagtg gcatctacag gtttgtgact ccaggagaac 4860
ggccctcggg catgttcgat tcttcggtcc tgtgtgagtg ctatgacgcg ggctgtgctt 4920
ggtatgagct cacgcccgct gagacctcgg ttaggttgcg ggcttaccta aatacaccag 4980
```

```
ggttgcccgt ctgccaggac catctggagt tctgggagag cgtcttcaca ggcctcaccc 5040
acatagatgc ccacttcctg tcccagacta aacaggcagg agacaacttt ccttacctgg 5100
tggcatatca agctacagtg tgcgccaggg ctcaagctcc accctccatcg tgggaccaaa 5160
tgtggaagtg tctcatacgg ctgaaaccta cactgcacgg gccaacaccc ctgctgtata 5220
ggctaggagc cgtccaaaat gaggtcatcc tcacacaccc cataactaaa tacatcatgg 5280
catgcatgtc ggctgacctg gaggtcgtca ctagcacctg ggtgctggta ggcggagtcc 5340
ttgcagcttt ggccgcatac tgcctgacga caggcagtgt ggtcattgtg ggcaggatca 5400
tcttgtccgg gaagccagct gtcgttcccg acagggaagt cctctaccag gagttcgatg 5460
agatggaaga gtgtgcctca caacttcctt acatcgagca gggaatgcag ctcgccgagc 5520
aattcaagca aaaggcgctc gggttgttgc aaacggccac caagcaagcg gaggctgctg 5580
ctcccgtggt ggagtccaag tggcgagccc ttgagacctt ctgggcgaag cacatgtgga 5640
atttcatcag cggaatacag tacctagcag gcttatccac tctgcctgga aaccccgcga 5700
tagcatcatt gatggcattt acagcttcta tcactagccc gctcaccacc caaaacaccc 5760
tcctgtttaa catcttgggg ggatgggtgg ctgcccaact cgctcctccc agcgctgcgt 5820
cagctttcgt gggcgccggc atcgccggag cggctgttgg cagcataggc cttgggaagg 5880
tgctcgtgga catcttggcg ggctatgggg caggggtagc cggcgcactc gtggcctta 5940
aggtcatgag cggcgaggtg ccctccaccg aggacctggt caacttactc cctgccatcc 6000
tctctcctgg tgccctggtc gtcggggtcg tgtgcgcagc aatactgcgt cggcacgtgg 6060
gcccgggaga gggggctgtg cagtggatga accggctgat agcgttcgct tcgcggggta 6120
accacgtctc ccctacgcac tatgtgcctg agagcgacgc tgcagcacgt gtcactcaga 6180
tcctctctag ccttaccatc actcaactgc tgaagcggct ccaccagtgg attaatgagg 6240
actgctctac gccatgctcc ggctcgtggc taagggatgt ttgggattgg atatgcacgg 6300
tgttgactga cttcaagacc tggctccagt ccaaactcct gccgcggtta ccgggagtcc 6360
ctttcctgtc atgccaacgc gggtacaagg gagtctggcg ggggacggc atcatgcaaa 6420
ccacctgccc atgcggagca cagatcgccg gacatgtcaa aaacggttcc atgaggatcg 6480
tagggcctag aacctgcagc aacacgtggc acggaacgtt ccccatcaac gcatacacca 6540
cgggaccttg cacaccctcc ccggcgccca actattccag ggcgctatgg cgggtggctg 6600
ctgaggagta cgtggaggtt acgcgtgtgg gggatttcca ctacgtgacg ggcatgacca 6660
ctgacaacgt aaagtgccca tgccaggttc cggccccga attcttcacg gaggtggatg 6720
gagtgcggtt gcacaggtac gctccggcgt gcaaacctct tctacgggag gacgtcacgt 6780
tccaggtcgg gctcaaccaa tacttggtcg ggtcgcagct cccatgcgag cccgaaccgg 6840
acgtaacagt gcttacttcc atgctcaccg atccctccca cattacagca gagacggcta 6900
agcgtaggct ggctagaggg tctcccccct ctttagccag ctcatcagct agccagttgt 6960
ctgcgccttc tttgaaggcg acatgcacta cccaccatga ctccccggac gctgacctca 7020
tcgaggccaa cctcttgtgg cggcaggaga tgggcggaaa catcactcgc gtggagtcag 7080
agaataaggt agtaattctg gactctttcg aaccgcttca cgcggagggg gatgagaggg 7140
agatatccgt cgcggcggag atcctgcgaa aatccaggaa gttcccctca gcgttgccca 7200
tatgggcacg cccggactac aatcctccac tgctagagtc ctggaaggac ccggactacg 7260
tccctccggt ggtacacgga tgcccattgc cacctaccaa ggctcctcca ataccacctc 7320
cacggagaaa gaggacggtt gtcctgacga aatccaatgt gtcttctgcc ttggcggagc 7380
tcgccactaa gaccttcggt agctccggat cgtcggccgt tgatagcggc acggcgaccg 7440
cccttcctga cctggcctcc gacgacggtg acaaaggatc cgacgttgag tcgtactcct 7500
ccatgccccc ccttgaaggg gagccggggg accccgatct cagcgacggg tcttggtcta 7560
ccgtgagtga ggaggctagt gaggatgtcg tctgctgctc aatgtcctat acgtggacag 7620
gcgccctgat cacgccatgc gctgcggagg aaagtaagct gcccatcaac ccgttgagca 7680
actctttgct gcgtcaccac aacatggtct acgccacaac atcccgcagc gcaagcctcc 7740
ggcagaagaa ggtcaccttt gacagattgc aagtcctgga tgatcattac cgggacgtac 7800
tcaaggagat gaaggcgaag gcgtccacag ttaaggctaa gcttctatct atagaggagg 7860
cctgcaagct gacgccccca cattcggcca aatccaaatt tggctatggg gcaaaggacg 7920
tccggaacct atccagcagg gccgttaacc acatccgctc cgtgtgggag gacttgctgg 7980
aagacactga aacaccaatt gacaccacca tcatgcaaa aagtgaggtt ttctcgtcc 8040
aaccagagaa gggaggccgc aagccagctc gccttatcgt attcccagac ctgggagttc 8100
gtgtatgcga gaagatggcc ctttacgacg tggtctccac ccttcctcag gccgtgatgg 8160
gctcctcata cggatttcaa tactccccca agcagcgggt cgagttcctg gtgaatacct 8220
ggaaatcaaa gaaatgccct atgggcttct catatgacac ccgctgtttt gactcaacgg 8280
tcactgagag tgacattcgt gttgaggagt caatttacca atgttgtgac ttggcccccg 8340
aggccagaca ggccataagg tcgctcacag agcggcttta catcggggggt ccctgacta 8400
actcaaaagg gcagaactgc ggttatcgcc ggtgccgcgc aagtggcgtg ctgacgacta 8460
gctgcggtaa tacccctcaca tgttacttga aggccactgc agcctgtcga gctgcaaagc 8520
tccaggactg cacgatgctc gtgaacggag acgacttggt cgttatctgt gaaagcgctgg 8580
gaacccagga ggatgcggcg ccctacgag ccttcacggg ggctatgact aggtattccg 8640
cccccccgg ggatccgccc caaccagaat acgacctgga gctgataaca tcatgttcct 8700
ccaatgtgtc agtcgcgcac gatgcatctg gcaaaagggt atactacctc acccgtgacc 8760
ccaccacccc ccttgcacgg gctgcgtggg agacagctag acacactcca atcaactctt 8820
ggctaggcaa tatcatcatg tatgcgccca ccctatgggc aaggatgatt ctgatgactc 8880
acttttttctc catccttcta gctcaagagc aacttgaaaa agccctggat tgtcagatct 8940
acggggcttg ctactccatt gagccacttg acctacctca gatcattgaa cgactccatg 9000
gtcttagcgc atttacactc cacagttact ctccaggtga gatcaatagg gtggcttcat 9060
```

```
gcctcaggaa acttggggta ccaccccttgc gaacctggag acatcgggcc agaagtgtcc 9120
gcgctaagct actgtcccag ggggggaggg ccgccacttg tggcagatac ctctttaact 9180
gggcagtaag gaccaagctt aaactcactc caatcccggc cgcgtcccag ctggacttgt 9240
ctggctggtt cgtcgctggt tacagcgggg gagacatata tcacagcctg tctcgtgccc 9300
gaccccgctg gtttccgttg tgcctactcc tactttctgt aggggtaggc atttacctgc 9360
tccccaaccg atgaacgggg agctaaccac tccaggcctt aagccatttc ctgttttttt 9420
tttttttttt tttttttttt tctttttttt tttctttcct ttccttcttt ttttcctttc 9480
tttttccctt ctttaatggt ggctccatct tagccctagt cacggctagc tgtgaaaggt 9540
ccgtgagccg catgactgca gagagtgctg atactggcct ctctgcagat catgt       9595
```

<210> 20
<211> 576
<212> DNA
<213> Hepatitis C virus

<400> 20
```
atgagcacca accccaagcc ccagcgcaag accaagcgga acaccaaccg gagaccccag 60
gacgtcaagt tcccaggagg aggccagatc gtgggcggcg tgtacctgct gccccgccgg 120
gggccccggc tgggcgtgcg cgccacccgc aagaccagcg agcgctccca gccaagaggc 180
agacgccagc cgatcccgaa ggcccgccgc cctgagggcc gggcttgggc ccagccaggc 240
tacccctggc ccctgtatgg caacgagggc ctgggatggg ctgggtggct cctcagcccc 300
cggggtcta ggcccagttg gggaccgacc gacccccgca ggcgcagccg caacctggga 360
aaggtgatcg acacgctcac ctgcggcttc gccgacttga tgggatacat ccctctggtg 420
ggggccctc tgggcggagc cgcgcgcgcc ctggctcacg gggtccgggt gctcgaggac 480
ggggtgaact acgccaccgg gaacctgccc ggctgcagct tctccatctt cctgctggcg 540
ctgctgagct gcctcaccat ccccgctagc gcatga                            576
```

<210> 21
<211> 1899
<212> DNA
<213> Hepatitis C virus

<400> 21
```
atggccccca tcaccgccta cagccagcag acccggggac tgctcggctg catcatcacc 60
tctctgacag gccgggataa gaaccaggtg gagggcgagg tgcaggtcgt ctcgaccgct 120
acccaaagct tcctggccac ctgtatcaac ggagtctgct ggacggtgta ccatggcgcc 180
ggcagcaaga ccctcgccgg gcctaagggc cccatcaccc agatgtacac caacgtggac 240
caggacctgg tgggctggca ggcgcccccc ggggcgagga gtatgcaccc atgcacctgc 300
gggagctctg acctgtatct ggtgaccaga catgccgatg tcatcccggt gaggcgtcgc 360
ggggacagta gagggagcct gctgagcccc cgccccgtca gctacctgaa ggggtccgtg 420
ggcggccccc tgctgtgccc ctctggccac gtggtcggca tcttcagggc cgccgtgtgc 480
acgcgcggcg tggccaaggc cgtggacttt atccccgtgg agagcatgga gaccaccatg 540
cgctcccccg tgttcaccga caacagcagc ccccccgccg tgcctcagac cttccaggtc 600
gcccacctcc atgctccgac gggctccggg aagtccacga aggtgcccgc cgcgtacgcg 660
gcccaggggat acaaggtgct ggtcctcaac cctagcgtgg ctgccacact cgggtttgga 720
gcgtacatga gcaaggcgca cggcatcgac cccaacatca gaactggcgt ccggaccatc 780
acaaccggcg ctcccatcac ttactctacc tacggcaagt tcctggctga tggggggtgt 840
agtgggggcg cgtacgatat tatcatctgc caggagtgcc actctaccga cagcacaca 900
atcctgggca tcggcaccgt cctcgaccag gctgagcag cgggcgcccg cctggtggtg 960
ctggccacgg ccactccccc cggctccgtc acggtgcccc accccaatat cgaggaggtg 1020
gccctgagca caacggcga gatcccattc tacggcaagg ctatcccgat cgaggcgatt 1080
aagggaggca gacatctgat cttctgccac agcaagaaga agtgcgacga gctcgccgcc 1140
aagctgagcg gcctcggact caacgccgtg gcttactaca ggggactgga cgtgtccgtg 1200
atcccgacca gcggagacgt ggtggtcgtg gccaccgacg ccctgatgac cggcttcacc 1260
ggagacttcg acagcgtcat cgactgcaac acctgcgtga cccagaccgt ggacttcagc 1320
ctggacccca ccttcaccat cgagaccacc acagtgcccc aggacgccgt gtcccgcagc 1380
cagcgccggg gccggacggg ccgcggccgg agtggcatct ataggttcgt gacccccgggc 1440
gagcgcccca gcggccatgtt cgatagttcc gtgctgtgcg agtgctacga cgccggatgc 1500
gcgtggtacg agctgacccc ggcggagacc tctgtccgcc tgagggctta cttgaatacc 1560
ccgggcctgc ccgtgtgcca ggatcatctc gagttctggg aatccgtctt caccggcctg 1620
acacacatcg acgcccattt cttgtcccaa accaagcagg ctggcgacaa tttcccgtat 1680
ctggtcgcgt accaggccac ggtgtgcgcg cgtgcgcagg ctccccccccc tagctgggat 1740
cagatgtgga agtgcctgat ccgcctgaag cccaccctgc atgggcccac ccccctgctg 1800
taccgcctgg gcgcggtgca gaacgaagtc accttgaccc accccatcac caagtacatc 1860
atggcgtgca tgtccgctga cctggaggtg gtcacctga                         1899
```

<210> 22
<211> 645
<212> DNA
<213> Hepatitis C virus

<400> 22
```
atgttttggg ccaagcatat gtggaacttc atcagcggca tccagtacct cgccgggctg 60
agcaccctcc cgggcaaccc cgcgatcgca agcctgatgg cgttcacagc gagcatcacc 120
tcccccctga ctacccagaa cacactgctg ttcaacatcc tggggggctg ggtcgccgct 180
cagctggccc ctccttccgc cgccagcgcc tttgtggggg cgggaatcgc cggggccgcc 240
gtcggctcca tcggactggg caaggtgctg gtcgacatcc tggcgggcta cggcgcggga 300
gtcgccggag ccctggtggc cttcaaggtg atgagcggag aggtgccaag cactgaggac 360
ctggtgaacc tgctgccgac gatcctgagc ccgggcgccc tggtggtggg cgtggtgtgt 420
gctgccatcc tcaggcgcca cgtgggcccg ggcgaggagg ccgtgcagtg gatgaaccgc 480
ctgatcgcct ttgcctcccg cggcaaccac gtcagcccta cacattacgt gcccgagagc 540
gatgccgccg cccgcgtgac ccagatcctg agctccctga ccatcaccca gctgctcaag 600
aggctgcacc agtggatcaa cgaggactgc tccacccctt gctga 645
```

<210> 23
<211> 1779
<212> DNA
<213> Hepatitis C virus

<400> 23
```
atgtccatgt cctacacctg gaccggcgcc ctgatcaccc cctgcgccgc cgaggagagc 60
aagctcccga ttaacccccct gtccaactct ctgctccgcc atcacaacat ggtgtatgcc 120
accacctccc gctctgcgag cctccgccag aagaaggtga cgttcgacag actgcaggtg 180
ctggacgacc attacaggga cgtgctgaag gaaatgaagg ccaaggctag caccgtgaag 240
gccaagctgc tcagcattga ggaggcttgc aagctgaccc cccccacag tgctaaatcc 300
aagttcggct acggcgccaa ggacgtgagg aacctgtcct cgcgcgctgt gaaccatatc 360
cgcagcgtgt gggaggacct gctcgaggac accgagaccc catcgcacac aaccatcatg 420
gccaagtccg aggtgttctg cgtgcagccg gagaaaggag gccgcaagcc agcccgcctg 480
atcgtcttcc ccgacctggg cgtgagagtc tgcgagaaga tggccctcta cgacgtggtg 540
tccaccctgc cgcaggccgt gatggggagt tcctacggct tccagtacag cccgaagcag 600
agggtggagt tcctggtgaa cacgtggaag tctaagaaat gccccatggg gttcagttac 660
ggaacaaggt gcttcgggag tactgtgacc gaatccgata tccgcgtgga ggagagcatc 720
taccagtgtt gtgacctcgc ccccgaggcg agacaggcca tccgctccct gaccgagagg 780
ctgtatatcg gcggcccact gaccaacagc aagggggcaga actgcggcta tcgccgttgt 840
cgggcctccg gggtgctcac cacctcttgc gggaacaccc tcacctgcta cctcaaggcg 900
accgctgcct gcagagccgc gaagctgcag gactgcacca tgctcgtgaa cggcgacgat 960
ctggtggtga tctgtgagtc cgcgggcacg caggaggacg cggcggccct gcgggcgttc 1020
acagaggcca tgacacgcta cagtgccccc cccggcgacc ccccccagcc cgaatacgat 1080
ctggagctca tcactagttg cagctcgaac gtgtctgtgg cccatgacgc ttctggcaaa 1140
cgggtgtatt atctgacgcg cgatcccacc accccccctcg ccagagccgc gtgggagaca 1200
gctcggcaca ccccctgtgaa ctcttggctg ggcaacatca tcatgtacgc ccctaccctg 1260
tgggctcgca tgatcctgat gacccacttc ttcagtatcc tcctcgctca ggagcagctg 1320
gagaaggcgc tcgactgcca gatctacggc gcctgctata gtatcgagcc tctcgacctg 1380
ccccagatca tcgagagact gcatgggctc agcgccttct ccctccatag ttactctcct 1440
ggagaaatta accgggtggc gagctgtctg cggaagctcg gcgtcccccc tctgcgcgtt 1500
tggcggcatc gcgccaggag tgtgagggcc aagctgctga gccaggcggg aagggccgcc 1560
acctgcggcc ggtatctctt caactgggcc gtgcgcacca agctcaagct caccccccatc 1620
cctgccgcca gtcagctgga tctcagtggg tggttcgtgg ccggctattc tggcggcgac 1680
atctaccact ccctcagcag ggcgcgcccc cgctggttcc ccctgtgcct gctgctcctg 1740
agcgtcggag tcggcatcta cctgctgccc aaccgctga 1779
```

<210> 24
<211> 3010
<212> PRT
<213> Hepatitis C virus

<400> 24
```
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
1               5                   10                  15
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25                  30
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
        35                  40                  45
Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
```

```
                    50                      55                      60
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
65                      70                      75                      80
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
                    85                      90                      95
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
                    100                     105                     110
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
            115                     120                     125
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
            130                     135                     140
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145                     150                     155                     160
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
                    165                     170                     175
Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr
                    180                     185                     190
Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys Ser
            195                     200                     205
Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Val Ile Met His Thr Pro
210                     215                     220
Gly Cys Val Pro Cys Val Gln Glu Gly Asn Ser Ser Arg Cys Trp Val
225                     230                     235                     240
Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr
                    245                     250                     255
Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Thr Ala Ala Phe Cys
                    260                     265                     270
Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Ile Phe Leu Val Ser
            275                     280                     285
Gln Leu Phe Thr Phe Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys
            290                     295                     300
Asn Cys Ser Ile Tyr Pro Gly His Val Ser Gly His Arg Met Ala Trp
305                     310                     315                     320
Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln
                    325                     330                     335
Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His
                    340                     345                     350
Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp
            355                     360                     365
Ala Lys Val Leu Ile Val Ala Leu Leu Phe Ala Gly Val Asp Gly Glu
370                     375                     380
Thr His Thr Thr Gly Arg Val Ala Gly His Thr Thr Ser Gly Phe Thr
385                     390                     395                     400
Ser Leu Phe Ser Ser Gly Ala Ser Gln Lys Ile Gln Leu Val Asn Thr
                    405                     410                     415
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
                    420                     425                     430
Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Ala His Lys Phe Asn
            435                     440                     445
Ser Ser Gly Cys Pro Glu Arg Met Ala Ser Cys Arg Pro Ile Asp Trp
    450                     455                     460
Phe Ala Gln Gly Trp Gly Pro Ile Thr Tyr Thr Lys Pro Asn Ser Ser
465                     470                     475                     480
Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Val
            485                     490                     495
Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser
            500                     505                     510
Pro Val Val Val Gly Thr Thr Asp Arg Ser Gly Val Pro Thr Tyr Ser
            515                     520                     525
Trp Gly Glu Asn Glu Thr Asp Val Met Leu Leu Asn Asn Thr Arg Pro
    530                     535                     540
Pro Gln Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe
545                     550                     555                     560
Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Val Gly Asn
                    565                     570                     575
Arg Thr Leu Ile Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala
            580                     585                     590
Thr Tyr Thr Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Leu
                            Page 11
```

```
                595                      600                     605
Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Leu Asn Phe
610                     615                     620
Ser Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Leu
625                     630                     635                     640
Asn Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asn Leu Glu Asp
                645                     650                     655
Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp
                660                     665                     670
Gln Ile Leu Pro Cys Ala Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly
                675                     680                     685
Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly
690                     695                     700
Val Gly Ser Ala Phe Val Ser Phe Ala Ile Lys Trp Glu Tyr Ile Leu
705                     710                     715                     720
Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp
                725                     730                     735
Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val
                740                     745                     750
Val Leu Asn Ala Ala Ser Val Ala Gly Ala His Gly Ile Leu Ser Phe
                755                     760                     765
Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Ala Pro
770                     775                     780
Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu
785                     790                     795                     800
Leu Ala Leu Pro Pro Arg Ala Tyr Ala Leu Asp Arg Glu Met Ala Ala
                805                     810                     815
Ser Cys Gly Gly Ala Val Leu Val Gly Leu Val Phe Leu Thr Leu Ser
                820                     825                     830
Pro Tyr Tyr Lys Val Phe Leu Thr Arg Leu Ile Trp Trp Leu Gln Tyr
                835                     840                     845
Phe Ile Thr Arg Ala Glu Ala His Met Gln Val Trp Val Pro Pro Leu
850                     855                     860
Asn Val Arg Gly Gly Arg Asp Ala Ile Ile Leu Leu Thr Cys Ala Val
865                     870                     875                     880
His Pro Glu Leu Ile Phe Asp Ile Thr Lys Leu Leu Leu Ala Ile Leu
                885                     890                     895
Gly Pro Leu Met Val Leu Gln Ala Gly Ile Thr Arg Val Pro Tyr Phe
                900                     905                     910
Val Arg Ala Gln Gly Leu Ile Arg Ala Cys Met Leu Val Arg Lys Val
                915                     920                     925
Ala Gly Gly His Tyr Val Gln Met Val Phe Met Lys Leu Gly Ala Leu
930                     935                     940
Thr Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala
945                     950                     955                     960
His Ala Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe
                965                     970                     975
Ser Ala Met Glu Thr Lys Val Ile Thr Trp Gly Ala Asp Thr Ala Ala
                980                     985                     990
Cys Gly Asp Ile Ile Leu Gly Leu Pro Val Ser Ala Arg Arg Gly Lys
                995                     1000                    1005
Glu Ile Phe Leu Gly Pro Ala Asp Ser Leu Glu Gly Gln Gly Trp Arg
                1010                    1015                    1020
Leu Leu Ala Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Val Leu
1025                    1030                    1035                    1040
Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu
                1045                    1050                    1055
Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr
                1060                    1065                    1070
Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys
                1075                    1080                    1085
Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val
                1090                    1095                    1100
Asp Leu Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Met
1105                    1110                    1115                    1120
Thr Pro Cys Ser Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His
                1125                    1130                    1135
Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu
                            Page 12
```

```
                1140                  1145                  1150
Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro
        1155                  1160                  1165
Leu Leu Cys Pro Ser Gly His Val Val Gly Val Phe Arg Ala Ala Val
    1170                  1175                  1180
Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser
1185                  1190                  1195                  1200
Met Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Thr Pro
                1205                  1210                  1215
Pro Ala Val Pro Gln Thr Phe Gln Val Ala His Leu His Ala Pro Thr
            1220                  1225                  1230
Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly
            1235                  1240                  1245
Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe
        1250                  1255                  1260
Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr
1265                  1270                  1275                  1280
Gly Val Arg Thr Ile Thr Thr Gly Gly Ser Ile Thr Tyr Ser Thr Tyr
            1285                  1290                  1295
Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile
            1300                  1305                  1310
Ile Ile Cys Asp Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu Gly
        1315                  1320                  1325
Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val
    1330                  1335                  1340
Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro
1345                  1350                  1355                  1360
Asn Ile Glu Glu Ile Gly Leu Ser Asn Asn Gly Glu Ile Pro Phe Tyr
            1365                  1370                  1375
Gly Lys Ala Ile Pro Ile Glu Ala Ile Lys Gly Gly Arg His Leu Ile
        1380                  1385                  1390
Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Thr
        1395                  1400                  1405
Gly Leu Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser
    1410                  1415                  1420
Val Ile Pro Pro Ile Gly Asp Val Val Val Val Ala Thr Asp Ala Leu
1425                  1430                  1435                  1440
Met Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr
            1445                  1450                  1455
Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile
        1460                  1465                  1470
Glu Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg
        1475                  1480                  1485
Gly Arg Thr Gly Arg Gly Arg Ser Gly Ile Tyr Arg Phe Val Thr Pro
    1490                  1495                  1500
Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys
1505                  1510                  1515                  1520
Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser
            1525                  1530                  1535
Val Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln
        1540                  1545                  1550
Asp His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile
        1555                  1560                  1565
Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro
    1570                  1575                  1580
Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro
1585                  1590                  1595                  1600
Pro Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro
            1605                  1610                  1615
Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln
        1620                  1625                  1630
Asn Glu Val Ile Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys
        1635                  1640                  1645
Met Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly
    1650                  1655                  1660
Gly Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Thr Thr Gly Ser Val
1665                  1670                  1675                  1680
Val Ile Val Gly Arg Ile Ile Leu Ser Gly Lys Pro Ala Val Val Pro
```

```
                    1685                1690                      1695
Asp Arg Glu Val Leu Tyr Gln Glu Phe Asp Glu Met Glu Glu Cys Ala
          1700                1705                1710
Ser Gln Leu Pro Tyr Ile Glu Gln Gly Met Gln Leu Ala Glu Gln Phe
          1715                1720                1725
Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Thr Lys Gln Ala Glu
          1730                1735                1740
Ala Ala Ala Pro Val Val Glu Ser Lys Trp Arg Ala Leu Glu Thr Phe
1745                1750                1755                1760
Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala
          1765                1770                1775
Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala
          1780                1785                1790
Phe Thr Ala Ser Ile Thr Ser Pro Leu Thr Thr Gln Asn Thr Leu Leu
          1795                1800                1805
Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Pro Pro Ser
          1810                1815                1820
Ala Ala Ser Ala Phe Val Gly Ala Gly Ile Ala Gly Ala Ala Val Gly
1825                1830                1835                1840
Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly
          1845                1850                1855
Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Val Met Ser Gly Glu
          1860                1865                1870
Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser
          1875                1880                1885
Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg
          1890                1895                1900
His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile
1905                1910                1915                1920
Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro
          1925                1930                1935
Glu Ser Asp Ala Ala Ala Arg Val Thr Gln Ile Leu Ser Ser Leu Thr
          1940                1945                1950
Ile Thr Gln Leu Leu Lys Arg Leu His Gln Trp Ile Asn Glu Asp Cys
          1955                1960                1965
Ser Thr Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Ile
1970                1975                1980
Cys Thr Val Leu Thr Asp Phe Lys Thr Trp Leu Gln Ser Lys Leu Leu
1985                1990                1995                2000
Pro Arg Leu Pro Gly Val Pro Phe Leu Ser Cys Gln Arg Gly Tyr Lys
          2005                2010                2015
Gly Val Trp Arg Gly Asp Gly Ile Met Gln Thr Thr Cys Pro Cys Gly
          2020                2025                2030
Ala Gln Ile Ala Gly His Val Lys Asn Gly Ser Met Arg Ile Val Gly
          2035                2040                2045
Pro Arg Thr Cys Ser Asn Thr Trp His Gly Thr Phe Pro Ile Asn Ala
2050                2055                2060
Tyr Thr Thr Gly Pro Cys Thr Pro Ser Pro Ala Pro Asn Tyr Ser Arg
2065                2070                2075                2080
Ala Leu Trp Arg Val Ala Ala Glu Glu Tyr Val Glu Val Thr Arg Val
          2085                2090                2095
Gly Asp Phe His Tyr Val Thr Gly Met Thr Thr Asp Asn Val Lys Cys
          2100                2105                2110
Pro Cys Gln Val Pro Ala Pro Glu Phe Phe Thr Glu Val Asp Gly Val
          2115                2120                2125
Arg Leu His Arg Tyr Ala Pro Ala Cys Lys Pro Leu Leu Arg Glu Asp
2130                2135                2140
Val Thr Phe Gln Val Gly Leu Asn Gln Tyr Leu Val Gly Ser Gln Leu
2145                2150                2155                2160
Pro Cys Glu Pro Glu Pro Asp Val Thr Val Leu Thr Ser Met Leu Thr
          2165                2170                2175
Asp Pro Ser His Ile Thr Ala Glu Thr Ala Lys Arg Arg Leu Ala Arg
          2180                2185                2190
Gly Ser Pro Pro Ser Leu Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala
          2195                2200                2205
Pro Ser Leu Lys Ala Thr Cys Thr Thr His His Asp Ser Pro Asp Ala
          2210                2215                2220
Asp Leu Ile Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn
```

```
2225                    2230                    2235                    2240
Ile Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe
                    2245                    2250                    2255
Glu Pro Leu His Ala Glu Gly Asp Glu Arg Glu Ile Ser Val Ala Ala
                2260                    2265                    2270
Glu Ile Leu Arg Lys Ser Arg Lys Phe Pro Ser Ala Leu Pro Ile Trp
            2275                    2280                    2285
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Leu Glu Ser Trp Lys Asp Pro
        2290                    2295                    2300
Asp Tyr Val Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Thr Lys
    2305                    2310                    2315                    2320
Ala Pro Pro Ile Pro Pro Pro Arg Arg Lys Arg Thr Val Val Leu Thr
                    2325                    2330                    2335
Glu Ser Asn Val Ser Ser Ala Leu Ala Glu Leu Ala Thr Lys Thr Phe
                2340                    2345                    2350
Gly Ser Ser Gly Ser Ser Ala Val Asp Ser Gly Thr Ala Thr Ala Leu
                2355                    2360                    2365
Pro Asp Leu Ala Ser Asp Asp Gly Asp Lys Gly Ser Asp Val Glu Ser
    2370                    2375                    2380
Tyr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu
    2385                    2390                    2395                    2400
Ser Asp Gly Ser Trp Ser Thr Val Ser Glu Glu Ala Ser Glu Asp Val
                2405                    2410                    2415
Val Cys Cys Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro
                2420                    2425                    2430
Cys Ala Ala Glu Glu Ser Lys Leu Pro Ile Asn Pro Leu Ser Asn Ser
            2435                    2440                    2445
Leu Leu Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala
    2450                    2455                    2460
Ser Leu Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp
    2465                    2470                    2475                    2480
Asp His Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr
                2485                    2490                    2495
Val Lys Ala Lys Leu Leu Ser Ile Glu Glu Ala Cys Lys Leu Thr Pro
                2500                    2505                    2510
Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg
            2515                    2520                    2525
Asn Leu Ser Ser Arg Ala Val Asn His Ile Arg Ser Val Trp Glu Asp
    2530                    2535                    2540
Leu Leu Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys
2545                    2550                    2555                    2560
Ser Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala
                2565                    2570                    2575
Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met
                2580                    2585                    2590
Ala Leu Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser
            2595                    2600                    2605
Ser Tyr Gly Phe Gln Tyr Ser Pro Lys Gln Arg Val Glu Phe Leu Val
    2610                    2615                    2620
Asn Thr Trp Lys Ser Lys Lys Cys Pro Met Gly Phe Ser Tyr Asp Thr
2625                    2630                    2635                    2640
Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Val Glu Glu
                2645                    2650                    2655
Ser Ile Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile
                2660                    2665                    2670
Arg Ser Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser
            2675                    2680                    2685
Lys Gly Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu
    2690                    2695                    2700
Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Thr Ala
2705                    2710                    2715                    2720
Ala Cys Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly
                2725                    2730                    2735
Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala
            2740                    2745                    2750
Ala Ala Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro
    2755                    2760                    2765
Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser
```

```
          2770                     2775                     2780
Cys Ser Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val
2785                     2790                     2795                     2800
Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp
                    2805                     2810                     2815
Glu Thr Ala Arg His Thr Pro Ile Asn Ser Trp Leu Gly Asn Ile Ile
                    2820                     2825                     2830
Met Tyr Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe
            2835                     2840                     2845
Phe Ser Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys
2850                     2855                     2860
Gln Ile Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln
2865                     2870                     2875                     2880
Ile Ile Glu Arg Leu His Gly Leu Ser Ala Phe Thr Leu His Ser Tyr
                    2885                     2890                     2895
Ser Pro Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly
                    2900                     2905                     2910
Val Pro Pro Leu Arg Thr Trp Arg His Arg Ala Arg Ser Val Arg Ala
            2915                     2920                     2925
Lys Leu Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Arg Tyr Leu
            2930                     2935                     2940
Phe Asn Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala
2945                     2950                     2955                     2960
Ala Ser Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly
            2965                     2970                     2975
Gly Asp Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Pro
            2980                     2985                     2990
Leu Cys Leu Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro
            2995                     3000                     3005
Asn Arg
    3010
```